

FINDING A SINGLE DEFECTIVE IN BINOMIAL GROUP-TESTING

by

Satindar Kumar

and Milton Sobel<sup>\*</sup>

Univ. of Wisc.--Milwaukee

Univ. of Minnesota

Milwaukee, Wisconsin

Minneapolis, Minnesota

Technical Report No. 139

June 1970

University of Minnesota

Minneapolis, Minnesota

<sup>\*</sup>This work was supported by the National Science Foundation Grant NSF-GP-11021 at the University of Minnesota.

# Abstract

The problem of finding a single defective item from an infinite binomial population is considered when the group-testing is possible, i.e., when we can test any number of units  $x$  simultaneously and find out if all  $x$  are good or if at least 1 of the  $x$  defective is present. An optimal procedure is obtained in the sense that it minimizes the expected number of tests required to find one defective. Upper and lower bounds are derived using information theory and the relation of our procedure to the Huffman algorithm and the corresponding cost is studied.

## 1. Introduction.

A group-test is a simultaneous test on a finite number  $x$  of units and we assume that each test has only two possible outcomes: i) either all the  $x$  units are satisfactory or ii) at least one of the  $x$  units is defective (we don't know which one or how many are defective). Using a binomial formulation we assume that each unit is defective with known probability  $p > 0$  and satisfactory with probability  $q = 1 - p$  and that the units are independent.

Our goal is simply to find a single defective unit by means of group-testing. The total number of units (or population size) is assumed to be countably infinite; this occurs, for example, in an assembly line production. This problem is related to the binomial group-testing problem discussed in [2] and [3] where all the units are classified. It is interesting to point out that although the procedure called  $R_1$  in [2] is not optimal for all  $q$ -values, a related procedure  $R'_{01}$  defined for the problem of this paper is optimal uniformly in  $q$ .

Regarding terminology, a defective set is a set of units known to contain at least one defective unit. If we have no such set at hand, we say that we have an H-situation.

In section 2 we define a procedure  $R'_{01}$  and we prove certain properties for it in section 3. Upper bounds on the expected number of tests  $F(m)$  required to "break up" a defective set of size  $m$  are obtained in section 4. Another procedure  $R'_{21}$  is introduced in section 5 which is shown to be equivalent to  $R'_{01}$  for  $q$  close to one. Lower bounds for any procedure that finds a single defective are derived in section 6. Section 7 contains a discussion on the optimality of the procedure  $R'_{01}$  and the relation of the breakup of the defective set under this procedure to the corresponding Huffman coding problem.

## 2. Procedure $R'_{01}$ for Known $q$ .

For any procedure  $R$  let  $E\{T|R\}$  and  $F\{m|R\}$ , respectively, denote the expected number of tests required to find one defective unit if we start with an H-situation and if we start with a defective

set of size  $m$ . If we start with a test on  $m$  units and use the fact that  $N$  is large (or infinite), then we obtain

$$E\{T|R\} = 1 + q^m E\{T|R\} + (1-q^m)F(m|R) \quad (1)$$

or, equivalently,

$$E\{T|R\} = \frac{1 + (1-q^m)F(m|R)}{1 - q^m} \quad (2)$$

For the particular procedure  $R'_{01}$  (in which we simply write  $E\{T\}$  and  $F(m)$ ), we choose the test group size  $m$  for any H-situation to be such that

$$E\{T\} = \min_{m=1,2,\dots} \left\{ \frac{1 + (1-q^m)F(m)}{1 - q^m} \right\} = \min_{m=1,2,\dots} \left\{ \frac{1 + pF^*(m)}{1 - q^m} \right\}, \quad (3)$$

where, by definition,

$$F^*(m) = \frac{(1-q^m)}{1-q} F(m). \quad (4)$$

If we start with a defective set of size  $m \geq 2$ , then the sample size  $x$  (to be taken exclusively from the defective set) is determined by

$$F(m) = 1 + \min_{1 \leq x < m} \left\{ \frac{q^x(1-q^{m-x})F(m-x) + (1-q^x)F(x)}{1 - q^m} \right\} \quad (5)$$

or, equivalently using  $F^*(m)$ ,

$$F^*(m) = \frac{1 - q^m}{1 - q} + \min_{1 \leq x \leq m} \{q^x F^*(m-x) + F^*(x)\} \quad (6)$$

the boundary conditions for this recursion are

$$F(1) = F^*(1) = 0 \quad \text{for all } q. \quad (7)$$

The values of  $F^*(m)$  were computed in Table IVA of [2] for

$m = 2(1)16$  for all  $q$ . For convenience let  $x = x(q)$  denote

the  $m$ -value that attains the minimum in (3) and let  $E^*\{T\} = (1-q^x)E\{T\}$ .

By direct computation we find that the expression in square brackets

in (3) for  $m = 1$  is less than that for  $m = 2$  when

$$1 - q - q^2 > 0 \quad (8)$$

or when  $q < (\sqrt{5} - 1)/2 = .618....$  Proceeding in a similar manner

we obtain the results given Table 1.

Thus, for example if  $q = .95$  then by Table I we take  $x = 14$

units for the first test group. If the test is successful they

are all good and we never use them again. Then we take another

group of size 14 and repeat the process. If a test is not successful

then by Table IVA of [2] we test 6 of the  $14$  units and proceed with  $F(8)$  if the 6 are all good or with  $F(6)$  if the 6 contain at least one defective. Using Table I, the value of  $E\{T\}$  required to find a defective is 5.761.

For any integer  $x$  let  $q_{x,x+1}$  denote the (unique) root in the unit interval of

$$1 - q^x - q^{x+1} = 0, \quad (9)$$

so that  $q_{0,1} = 0$ ,  $q_{1,2} = .618\dots$ , etc. Then we make the

Conjecture:

For all  $q$  in the interval  $[q_{x-1,x}, q_{x,x+1}]$  the integer  $x$  which is the root of (9) achieves the minimum in (3) and the value of  $E^*\{T\}$  in this interval is given by

$$E^*\{T\} = (1+\alpha)(1-q^x) + q^{x-2\beta}, \quad (10)$$

where  $\alpha = \alpha(x)$  and  $\beta = \beta(x)$  are uniquely defined by

$$x = 2^\alpha + \beta \quad (0 \leq \beta < 2^\alpha) \quad (11)$$

This conjecture is essentially a conjecture that the last dividing point for  $F_1^*(x)$  before  $q = 1$  (see Table IV A in [2]) is less than the corresponding dividing point  $q_{x-1,x}$  between  $x - 1$  and  $x$ ; these  $q$ -values are given in Table VII (H-situation) in [2]. If this inequality holds then we can use equation (23) in [2] (given also in (5.5) below) to show that for  $q$  in the interval  $[q_{x-1,x}, q_{x,x+1}]$  the minimum in (3) above is attained at  $m = x$ ; this is carried out in section 3. Furthermore we can then substitute (23) of [2] into (3) and we easily obtain (10). This conjecture has been verified by the computations in Table I for  $x = 2(1)15$ .

### 3. A Property of $F(m)$ .

In this section we wish to show that if we start with  $F(m)$  the next group test size  $x$  will always be at most  $m/2$ ; this is related to the discussion in Section VI and in particular to (21) in [2]. To show this we first prove with the help of lemma 1 below that  $F(m)$  is a nondecreasing function of  $m$  for any  $q$ .



We first note from (2.4) that the inequality  $F(m) \leq F(m+1)$  is equivalent to

$$F^*(m) \leq F^*(m+1) - q^m F(m). \quad (12)$$

This is proved by noting that

$$0 \leq \sum_{i=1}^{m+1} q^{i-1} \{F(m+1) - F(m)\} = F^*(m+1) - F^*(m) - q^m F(m). \quad (13)$$

For the  $F(m+1)$ -situation we use  $y$  to denote a possible size for the next group test.

Lemma 1:

If  $F(1) \leq F(2) \leq \dots \leq F(m)$  and  $y \leq (m+1)/2$  then, for the  $F(m+1)$ -situation,  $y$  is preferable to  $m+1-y$ .

Proof:

Consider the quantity in braces on the right side of (5) with  $x = y$  and  $x = m+1-y$  and denote these by  $F_1$  and  $F_2$ , respectively. Then after algebraic simplification we obtain

$$F_2 - F_1 = \frac{(1-q^y)(1-q^{m+1-y})}{1-q^{m+1}} (F(m+1-y) - F(y)). \quad (14)$$

Since  $1 \leq y \leq (m+1)/2$  it follows from the hypothesis that the last factor in (14) is nonnegative and this proves lemma 1.

Theorem 1:

For all  $q$  and any  $m \geq 2$

$$F(m) \leq F(m+1). \quad (15)$$

Proof:

Since  $0 = F(1) < F(2) = 1$ , it suffices to show by induction that if  $F(1) \leq F(2) \leq \dots \leq F(m)$  then  $F(m) \leq F(m+1)$ . From (4) and (5) we have

$$F^*(m) = \frac{1 - q^m}{p} + \min_{1 \leq x \leq m/2} \{q^x F^*(m-x) + F^*(x)\} \quad (16)$$

since by lemma 1 we can restrict  $x$  to at most  $m/2$ . Using (12)

with  $m$  replaced by  $m - x$ , we replace  $F^*(m-x)$  by an upper

bound and obtain

$$\begin{aligned} F^*(m) &\leq \frac{1 - q^{m+1}}{p} - q^m F(m) + \min_{1 \leq x \leq m/2} \{q^x F^*(m+1-x) + F^*(x) \\ &\quad + q^m \{F(m) - F(m-x) - 1\}\} \end{aligned} \quad (17)$$

We now consider separately the cases for  $m$  odd and  $m$  even.

Case 1:  $m = 2n$

We note that  $\lceil \frac{m}{2} \rceil = \lceil \frac{m+1}{2} \rceil$  where  $[x]$  is the largest integer  $\leq x$ . Furthermore, since  $1 \leq x \leq \frac{m}{2}$  we have  $x \leq 2n - x$  and, by hypothesis,  $F(x) \leq F(2n-x)$ . Then by (5) for any  $x$  with  $1 \leq x \leq n$

$$F(2n) \leq 1 + PF(x) + (1-P)F(2n-x) \leq 1 + F(2n-x), \quad (18)$$

where  $0 \leq P \leq 1$ . Hence the expression in braces in (17) is nonpositive and can be dropped. This gives the inequality

$$\begin{aligned} F^*(m) &\leq \frac{1 - q^{m+1}}{p} + \min_{1 \leq x \leq \frac{m+1}{2}} \{q^x F^*(m+1-x) + F^*(x)\} - q^m F(m) \\ &= F^*(m+1) - q^m F(m) \end{aligned} \quad (19)$$

which, by (12), is what we need to prove (15).

Case 2:  $m = 2n + 1$

Clearly we can write  $x \leq (m+1)/2$  under the Min sign in (17) since the minimum is attained for some  $x \leq m/2$ . If the expression in braces in (17) is nonpositive, then the proof is

the same as Case 1. We now consider 2 subcases

Case 2A:  $F(2n+1) - F(n) - 1 > 0$  and for  $F(2n+2)$  the optimal integer  $x \leq n$ . Since  $n + 1$  does not minimize  $q^x F^*(n+1-x) + F^*(x)$  and the expression in braces in (17) is positive, it follows that the minimum in (17) must be attained for some  $x \leq n$ . For  $x \leq n$ , by (18) the expression in braces is nonpositive and the same proof goes through.

Case 2B:  $F(2n+1) - F(n) - 1 > 0$  and for  $F(2n+2)$  the optimal integer  $x = n + 1$ . Clearly

$$F(2n+2) = 1 + F(n+1) \quad (20)$$

Moreover for some  $P(0 \leq P \leq 1)$  if we take  $x = n$  and use the hypothesis

$$F(2n+1) \leq 1 + PF(n+1) + (1-P)F(n) \leq 1 + F(n+1). \quad (21)$$

From (20) and (21) we get the desired result

$$F(2n+1) \leq F(2n+2), \quad (22)$$

which completes the proof of the theorem.

Corollary:

Starting with any defective set of size  $m \geq 2$ , the optimal  $x$  for the size of the next group test is at most  $\lfloor m/2 \rfloor$ .

Since the hypothesis of the lemma above is now proved, the conclusion that  $x$  is preferable to  $m - x$  is equivalent to this corollary.

4. Upper Bounds on  $F(m)$  under Procedure  $R'_{01}$ .

In this section we describe useful bounds that hold only for procedure  $R'_{01}$ ; later in section 6 we describe general lower bounds on  $F(m)$  that hold for any group-testing procedure.

Lemma 2:

If  $\gamma = \gamma(m)$  is defined by  $2^{\gamma-1} < m \leq 2^\gamma$ , then

$$F(m) \leq \gamma. \quad (23)$$

Proof:

Add "fictitious" good units to the "real" units so that the total is  $2^\gamma$ . Then, by taking  $x = m/2 = 2^{\gamma-1}$  in (5), we obtain for  $m = 2^\gamma$

$$F(2^\gamma) \leq 1 + F(2^{\gamma-1}). \quad (24)$$

Repeating this inequality and using the fact that  $F(2) = 1$ ,

we obtain

$$F(2^\gamma) \leq \gamma - 1 + F(2) = \gamma. \quad (25)$$

We impose the condition that no test should be carried out on fictitious units alone. Since this can only reduce the number of tests, it follows that  $\gamma$  is still an upper bound and this proves the lemma.

The above bound does not depend on  $q$ . It is possible to obtain an improved upper bound on  $F(m)$  that depends on  $q$ .

For this purpose we write  $m$  in its binary expansion form

$$m = 2^{r_1} + 2^{r_2} + \dots + 2^{r_s}, \quad (26)$$

where  $r_1 > r_2 > \dots > r_s \geq 0$  and  $s > 0$  are integers. Let

$y_1 = 2^{r_1}$  and  $x_1 = m - y_1$ ; let  $y_2 = 2^{r_2}$  and  $x_2 = x_1 - y_2$ , etc.;

at the last step  $y_s = 2^{r_s}$  and  $x_s = 0$ . Using the right side

of (5) with  $x = x_1$  instead of taking the minimum, we obtain an

upper bound for  $F(m)$ . Similarly with  $m$  replaced by  $m - y_1$ ,

we use the right side of (5) with  $x = x_2$  to get an upper bound

for  $F(m - y_1)$ . Repeating this and using Lemma 2, we obtain

$$\begin{aligned}
 F(m) &\leq \frac{q^{x_1}(1-q^{y_1})}{1-q^m} F(y_1) + \frac{1-q^{x_1}}{1-q^m} F(x_1) \\
 &\leq 1 + r_1 \frac{q^{x_1}(1-q^{y_1})}{1-q^m} + \frac{(1-q^{x_1})}{1-q^m} \left\{ 1 + \frac{q^{x_2}(1-q^{y_2})}{1-q^{x_1}} F(y_2) \right. \\
 &\quad \left. + \frac{1-q^{x_2}}{1-q^{x_1}} F(x_2) \right\} \\
 &\leq 1 + \frac{1}{1-q^m} \sum_{i=1}^s \{1 - q^{x_i} + r_i q^{x_i}(1-q^{y_i})\}.
 \end{aligned} \tag{27}$$

For example, if  $m = 21$  we have  $21 = 2^4 + 2^2 + 2^0$  so that

$s = 3$ ,  $y_1 = 16$ ,  $x_1 = 5$ ,  $y_2 = 4$ ,  $x_2 = 1$ ,  $y_3 = 1$  and  $x_3 = 0$ . Then

the last term vanishes and we obtain

$$F(21) \leq 3 + \frac{1}{1-q^{21}} \{q(1-q^{20}) + q^5(1-q^{16})\}. \tag{28}$$

For  $q \rightarrow 1$  this leads to the upper bound  $4 + \frac{5}{7}$  for all  $q$ ,

which is less than the result  $\gamma = 5$  obtained by Lemma 4.1.

##### 5. Procedure $R'_{21}$ .

In this section we discuss an alternate procedure  $R'_{21}$  for

the H-situation that is based on information theory; this procedure is used in [2] and [3] as an alternate procedure for classifying all the units in a binomial sample. When starting with a defective set of any size  $m \geq 2$  the new procedure  $R'_{21}$  is defined exactly the same as procedure  $R'_{01}$ . By the computation in Table I and the theorem below, it follows that these 2 procedures are identical for  $q < .9563$  (corresponding to  $x \leq 15$ ) and also for  $q$  sufficiently close to one; it is conjectured that they are identical for all  $q$ , but this has not been proved.

For the H-situation we take a sample of size  $x$  where  $x$  is the integer that maximizes

$$-\{q^x \log_2 q^x + (1-q^x) \log_2 (1-q^x)\}. \quad (29)$$

The dividing point  $q_{x,x+1}$  between  $x$  and  $x+1$  is shown in [2] to be the unique real root (in the unit interval) of

$$1 - q^x - q^{x+1} = 0. \quad (30)$$

For the case of a defective set of size  $m$  the recursion (5)

and the boundary condition (7) are again used.



Theorem 2:

For  $q$  close to 1, the procedures  $R'_{21}$  and  $R'_{01}$  are equivalent.

Proof:

We need only prove the equivalence for the H-situation since the two procedures are defined to be equivalent for the case of a defective set of size  $m \geq 2$ . For procedure  $R'_{01}$  the dividing point between  $x$  and  $x + 1$  is determined by the root of

$$\frac{1 + pF^*(x)}{1 - q^x} = \frac{1 + pF^*(x+1)}{1 - q^{x+1}} \quad (31)$$

or, equivalently,

$$q^x = (1 - q^x)F^*(x+1) - (1 - q^{x+1})F^*(x). \quad (32)$$

For  $q$  close to 1 we have by (23) of [2] that

$$pF^*(x) = \alpha(1 - q^x) + q^{x-2\beta}(1 - q^{2\beta}), \quad (33)$$

where  $\alpha = \alpha(x)$  and  $\beta = \beta(x)$  are defined in (11). Using this

in (32), we obtain

$$pq^x = (1-q^x)(1-q^{x+1})\{\alpha(x+1) - \alpha(x)\} + (1-q^x)\{q^{x+1-2\beta(x+1)} - q^{x+1}\} \\ - (1-q^{x+1})\{q^{x-2\beta(x)} - q^x\} . \quad (34)$$

To simplify this we consider two cases according as  $x + 1$  is not or is a power of 2.

Case 1:  $\alpha(x+1) = \alpha(x)$

Then  $\beta(x+1) = 1 + \beta(x)$  and (34) becomes

$$pq^x = (1-q^x)(q^{x-2\beta(x)-1} - q^{x+1}) - (1-q^{x+1})(q^{x-2\beta(x)} - q^x) \quad (35)$$

and, after simplification, this reduces to  $1 - q^x - q^{x+1} = 0$ ,

which gives rise to exactly the same dividing point between  $x$

and  $x + 1$  as for procedure  $R'_{21}$ .

Case 2:  $\alpha(x+1) = 1 + \alpha(x)$

Then  $\beta(x) = 2^{\alpha(x)} - 1$ ,  $\beta(x+1) = 0$  and hence

$$x - 2\beta(x) = 2^{\alpha(x+1)} - 1 - 2\{2^{\alpha(x)} - 1\} = 1. \quad (36)$$

Hence (34) becomes

$$pq^x = (1-q^x)(1-q^{x+1}) - (1-q^{x+1})(q - q^x) = p(1-q^{x+1}) \quad (37)$$

and this again reduces to  $1 - q^x - q^{x+1} = 0$ . Since the dividing point is the same as for procedure  $R'_{21}$ , the procedures are identical for  $q$  sufficiently close to 1.

6. Lower Bounds on  $F(m|R)$  and  $E\{T|R\}$  for any procedure  $R$ .

A lower bound on  $F(m|R)$  for any procedure that accomplishes the goal of finding a single defective is obtained from information theory. The total entropy associated with a defective set of size  $m$  is

$$\sum_{i=1}^m \frac{pq^{i-1}}{1 - q^m} \log_2 \left( \frac{1 - q^m}{pq^{i-1}} \right). \quad (38)$$

This entropy (or uncertainty) is reduced to zero in  $F(m|R)$  tests.

Since the maximum reduction per test is one, it follows that

$$F(m|R) \geq \sum_{i=1}^m \frac{pq^{i-1}}{1 - q^m} \log_2 \left( \frac{1 - q^m}{pq^{i-1}} \right) \quad (39)$$

and after algebraic simplification

$$F(m|R) \geq \frac{1}{p} I(p) - \frac{1}{1 - q^m} I(1 - q^m), \quad (40)$$

where  $I(p) = p \log_2 \frac{1}{p} + q \log_2 \frac{1}{q}$ . It is easily shown by

differentiation that  $I(p)/p$  is strictly decreasing in  $p$  and, since  $p > 1 - q^m$ , the lower bound in (40) is positive.

We can also obtain a lower bound on  $E\{T|R\}$  for any procedure that finds a single defective; we actually obtain 2 lower bounds and show that the one that makes use of (40) and is based on procedure  $R'_{01}$  is the better one.

Using the same argument as above with the disjoint, exhaustive set of probabilities  $pq^{i-1}$  ( $i = 1, 2, \dots$ ) we easily obtain

$$E\{T|R\} \geq - \sum_{i=1}^{\infty} pq^{i-1} \log_2 pq^{i-1} = \frac{1}{p} I(p) . \quad (41)$$

The result (40) can be applied to  $F(m)$  in (3) to yield another lower bound on  $E\{T\}$  for procedure  $R'_{01}$ . We obtain

$$\begin{aligned} E\{T\} &= \min_{m=1,2,\dots} \left\{ \frac{1}{1 - q^m} + F(m) \right\} \\ &\geq \frac{1}{p} I(p) + \min_{m=1,2,\dots} \left\{ \frac{1 + I(1-q^m)}{1 - q^m} \right\} . \end{aligned} \quad (42)$$

Since  $I(p)$  has a maximum of one, it follows that  $I(1-q^m) \leq 1$  and hence the lower bound in (42) is at least as large as that in (41). Since these lower bounds are generally not attainable,

it should not be assumed that there exist procedures  $R$  with  $E\{T|R\}$  between the right members of (41) and (42). In fact the construction of the procedure  $R'_{01}$  leads us to assert that the right side of (42) is also a lower bound for any procedure  $R$  that finds a single defective; this point is also discussed below in connection with optimality properties.

#### 7. Optimality Discussion.

For the case of the so-called  $F(m)$ -situation we have used the principle of "backward optimization" to define the procedure  $R'_{01}$ . For this subproblem with the given value of  $q$  this procedure  $R'_{01}$  is therefore optimal regardless of the value of  $m$  that we start with. Thus we have a cost function  $E\{T|m\}$  for each value of  $m$ . For the  $H$ -situation what we do under procedure  $R'_{01}$  is simply to find the  $m$ -value that minimizes this cost function. Thus  $R'_{01}$  is an optimal procedure for the overall problem of finding a single defective. This explains why the right side of (42) must be an improved lower bound for all procedures that find a single defective.

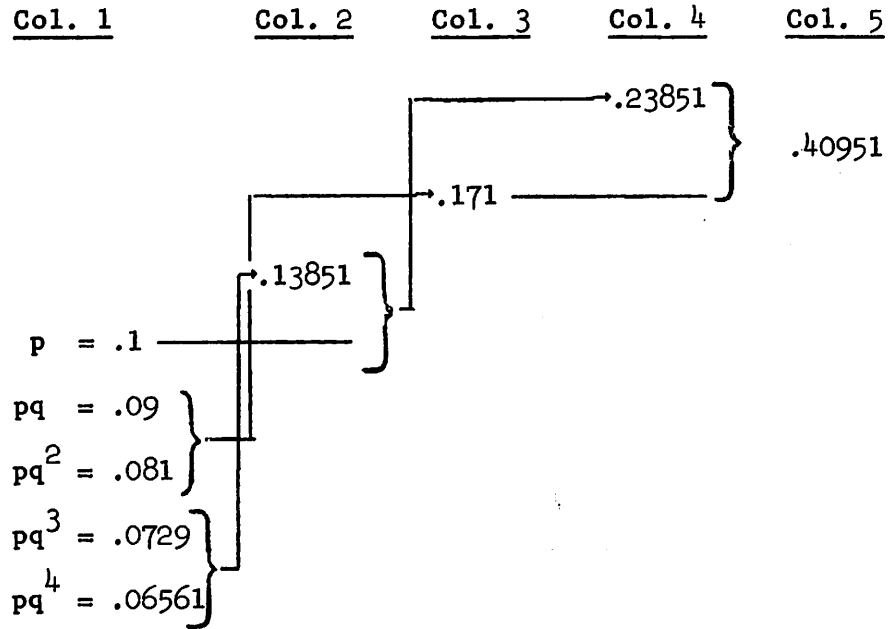
In this connection we should point out that another lower bound for  $F(m)$  is obtained from a consideration of the cost of the Huffman code when we have  $m$  states of nature with probabilities proportional to  $pq^{i-1}$  ( $i = 1, 2, \dots, m$ ). The relation of the Huffman code to the group-testing procedure is described in [3] (see section 14) and the Huffman algorithm is given in [1]. The discussion in [3] indicates that for  $F(m)$  the Huffman algorithm of adding the two smallest probabilities, reordering the resulting set, adding the two smallest again, reordering again, etc. will yield exactly the same  $F(m)$ -values as we obtain for the procedure  $R'_{01}$ . This result has been verified for small values of  $m$ , but has not been proved in general. If true, it provides an alternate way of computing  $F(m)$  for procedure  $R'_{01}$  for any particular values of  $q$  and for all values of  $m$ ; for  $m = 2(1)16$  the results are given in [2]. The details for obtaining the value of  $F(m)$ , which is the same as the Huffman cost, are described at the end of section 12 in [3] and we need only give a brief illustration of this. Suppose, for example  $m = 5$  and  $q = .9$  so that the

5 states of nature have probabilities proportional to

(43)

$$\{pq^{i-1} (i = 1, 2, 3, 4, 5)\} = \{.1, .09, .081, .0729, .06561\}$$

The Huffman algorithm yields



The Huffman cost is obtained by summing the four numbers

appearing in columns 2 through 5 and dividing by  $1 - q^5 = .40951$ ;

this gives a Huffman cost of  $.95753/.40951 = 2.33823$ . The

polynomial ratio corresponding to this calculation is

$$\frac{q^3(1-q^2) + q(1-q^2) + \{q^3(1-q^2)+p\} + \{q^3(1-q^2)+q(1-q^2)+p\}}{1 - q^5}$$

$$= p(2 + 2q + 2q^2 + 3q^3 + 3q^4)/(1-q^5), \quad (44)$$

which agrees with the result for  $F(m)$  for  $q = .90$  obtainable from Table IV A of [2]. Since the Huffman cost represents a lower bound for the expected number of tests, this shows that our procedure for breaking up the set of 5 defectives is optimal.



Table I

<u>x</u>	<u>E*(T)</u>	<u>for q-values in</u> <u>the interval</u>
1	1	[0, .6180]
2	$2 - q^2$	[.6180, .7549]
3	$2 + q - 2q^3$	[.7549, .8192]
4	$3 - 2q^4$	[.8192, .8567]
5	$3 + q^3 - 3q^5$	[.8567, .8813]
6	$3 + q^2 - 3q^6$	[.8813, .8987]
7	$3 + q - 3q^7$	[.8987, .9116]
8	$4 - 3q^8$	[.9116, .9216]
9	$4 + q^7 - 4q^9$	[.9216, .9296]
10	$4 + q^6 - 4q^{10}$	[.9296, .9361]
11	$4 + q^5 - 4q^{11}$	[.9361, .9415]
12	$4 + q^4 - 4q^{12}$	[.9415, .9460]
13	$4 + q^3 - 4q^{13}$	[.9460, .9499]
14	$4 + q^2 - 4q^{14}$	[.9499, .9533]
15	$4 + q - 4q^{15}$	[.9533, .9563]

Remark:

Under the conjecture in Section 2 the general result for all

x is

$$E^*(T) = (1+\alpha)(1-q^x) + q^{x-2\beta}$$

and the appropriate dividing points for these polynomials are

given in Table VII of [2] for  $x = 1(1)100$ .

### References

- [1] Huffman, David A. (1952). A method for the construction of minimum redundancy codes. Proc. I.R.E. 40, 1098-1101.
- [2] Sobel, M. and Groll, P. A. (1959). Group-testing to eliminate efficiently all defectives in a binomial sample. Bell System Tech. Jour. 38, 1179-1252.
- [3] Sobel, M. (1960). Group-testing to classify efficiently all defectives in a binomial sample. Information and Decision Processes, edited by R. E. Machol, McGraw-Hill, 127-161.